



# The sound of quotation marks: Prosodic characteristics of subclausal quotation in English

Bethany Sturman

University of California, Los Angeles, United States

bsturman@ucla.edu

## Abstract

Subclausal quotation (SQ, aka partial quotation or mixed quotation) is a phenomenon in which the utterance contains material, typically an NP or DP, attributed to a source other than the speaker (e.g. *Angela thinks it's an insult when she calls me a % "genius."*). Though there has been significant discussion of SQ from a semantic perspective, little work has been done to understand the intonation accompanying this construction. Using data from a production experiment, National Public Radio, and television comedies such as “The Office,” this paper seeks to identify the fundamental prosodic features that characterize SQ intonation in Mainstream American English using the MAE\_ToBI system. These features include an emphatic juncture (i.e. an IP juncture with a plateau boundary tone sequence followed by an obligatory pause) delimiting the start of the quotation, a pitch range reset (often an expansion) on the quoted material, and an IP break marking the end of the quotation. This break is realized with an L-L% or L-H% boundary tone sequence. Following the quotation, the speaker returns to their pre-quotation pitch register. SQ can optionally be marked lexically using “quote” to indicate the beginning, although the accompanying “unquote” is rarely employed.

**Index Terms:** quotation, subclausal quotation, semantics, intonation-semantics interface, emphatic juncture

## 1. Introduction

Subclausal quotation (SQ, aka partial quotation or mixed quotation) is a phenomenon in which the utterance contains material attributed to a source other than the speaker (e.g. *What if she brought her % “partner” % into work?*). This construction is produced with a highly salient intonational contour and is often accompanied by other paralinguistic features like finger quotations and eyebrow wagging. Though there has been significant discussion of SQ from a semantic perspective, little work has been done to understand the intonation accompanying this construction in English (see [1] for related work in German). Adopting the Autosegmental-metrical (AM) model of intonational phonology for English and the MAE\_ToBI system [2, 3, 4, 5], this paper seeks to identify the fundamental prosodic features that characterize SQ at a phonological level. One of the primary motivations for identifying these core phonological markers is to establish a tenable starting place to explore how we might integrate intonation more rigorously into the formal semantic analysis of a specific phenomenon, with the ultimate aim of moving toward the inclusion of intonation in formal semantics more broadly.

## 2. Background

### 2.1. Subclausal quotation versus full clause quotation

In SQ, the size of the quoted material is typically a relatively small syntactic constituent, such as NP or DP. The amount of quoted material is one of the primary distinguishing features between subclausal quotation and full clause quotation. Additionally, full clause quotation is generally used to deliver a direct speech report with a specified source for the quoted material [6] (e.g. *Juilo said, “I passed the exam with flying colors.”*).

Previous work has examined the prosodic markers of full clause quotation, finding an intonation phrase break preceding the quoted material as well as an expanded pitch range within the quotation [7]. It is still an unexplored question, however, what prosodic markers speakers use to signal instances of SQ.

### 2.2. Semantic uses of subclausal quotation

In contrast to the semantic usage of full clause quotation, SQ need not have an explicitly specified source, and SQ can be used in a much broader set of semantic and pragmatic contexts [8, 9, 10]. For example, SQ can be used to mimic a speaker’s use of or pronunciation of a particular word (e.g. *In California, Beth carries her groceries in “b[æ]gs,” but in Minnesota she carries them in “b[e]gs.”*). SQ can also be used when uttering a predicate of personal taste (e.g. adjectives like *tasty* or *fun* [11]) which the speaker would not endorse in the context (e.g. *Miriam thought the obstacle course was “super easy,” but I barely made it past the monkey bars.*). Relatedly, SQ can be used to indicate perspective shift [9, 12]. Perspective shift is a semantic strategy a speaker can employ when they wish to use a particular phrase that they do not want attributed to their set of beliefs within a larger utterance that they do endorse (e.g. *He said Ukraine should turn to its “good old neighbor” Russia.*).

## 3. Materials

The data used in this analysis come from several sources, including National Public Radio (NPR) broadcast recordings (both news reporting and interviews with more spontaneous speech), television shows such as “The Office” and “Barry,” and a production task that was a component of a paired production-perception experiment.

### 3.1. NPR and television shows

The television show tokens were identified by the actors producing the SQ in conjunction with using air quotes. The NPR tokens were identified through either 1) the presence of orthographic quotation marks marking SQ in the transcript (though titles marked with quotes were not analyzed) or 2) the perception of SQ by the author in a semantically reasonable context for SQ. These tokens were then confirmed to be SQ by native

English-speaking colleagues. To mitigate the potential circularity introduced the second data collection method, future work will include an experimental task in which linguistically naive native English speakers provide an orthographic transcription of each of these tokens. If participants' transcriptions consistently include orthographic quotes, such results would suggest that the intonation identified by the author in these tokens is indeed the prosody canonically associated with SQ.

### 3.2. Production-perception experiment

The production data come from a paired production-perception experiment investigating the prosody of SQ and its perspective shifting properties. Specifically, in the condition of the experiment relevant for this paper, the talker participant was shown a sentence that contained SQ surrounding the object NP (e.g. *Larry challenged an “alarming rule” at the board meeting.*), chose how they interpreted the sentence (e.g. *Who considered the rule alarming?* [Only Larry, everyone]), then read the sentence aloud with the intent of communicating their chosen interpretation to the listener participant. This paradigm was successful in eliciting emotionally engaged speech for many of the participant pairs. Six of the target items in the experiment were from this condition, and there were 72 total sentences in the experiment. The 30 tokens from the SQ experiment analyzed in this paper were produced by 21 native English speakers and represent a subset of the total data collected due to time constraints.

### 3.3. Data analysis

All of the data were transcribed in the MAE-ToBI system by the author, and these transcriptions were checked by other expert annotators. There were 61 total tokens (24 from NPR, 7 from the television shows, and 30 from the experiment). The average number of words contained in the SQ was  $2.49 \pm .19$ . Of the 61 tokens, 41 (67%) occurred sentence medially.

## 4. The prosodic markers of SQ

Unlike some other perspective shifting phenomena, such as transparent free relatives (e.g. *Allen poured what he called a beergarita.*) or even full clause quotation, SQ need not be marked with any lexical material. (The word “quote” can be used, but this is optional and will be discussed in more detail later.) As such, the intonation associated with SQ is doing the bulk of the heavy lifting to communicate the presence of SQ and the speaker’s intended meaning.

It was found that the characterization of the prosody of SQ consists of three fundamental intonational features that can be represented phonologically. Specifically, the start of the quotation is marked with an emphatic juncture; there is a pitch range reset (and often an expansion) on the quoted material, and following the quotation, the speaker returns to their pre-quotation register; and an IP break marks the end of the quotation. If the quotation is sentence medial, the boundary tone sequence of the quotation final IP is typically L-H% or L-L%. These features were consistently found across all three data sources. Each will be explored in more detail in the sections below.

### 4.1. Initial emphatic juncture

#### 4.1.1. Emphatic juncture background

The Emphatic Juncture [13] is a specialized sub-type of the IP boundary that functions to highlight the material following

the juncture as prominent or to instantiate a perspective shifting phenomenon such as transparent free relatives [14, 15]. Acoustically, the Emphatic Juncture is characterized by IP-level final lengthening, a plateau boundary tone (H-L% or !H-L%), and an obligatory pause. This pause is intentionally planned by the speaker rather than a disfluency. Some examples of an emphatic juncture marking the start of perspective shift in transparent free relatives can be seen in Figs. 1 and 2. The Emphatic Juncture is labeled on the Breaks tier using 4e and is highlighted with a box in the following examples.

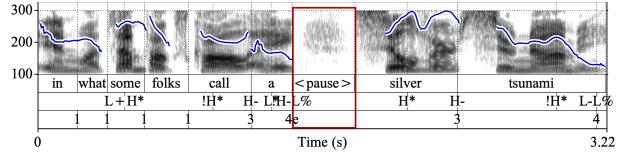


Figure 1: Pitch track of ‘In what some folks call a % silver tsunami.’ (data from NPR) This example demonstrates the presence of an emphatic juncture preceding the perspective-shifted material in a transparent free relative.

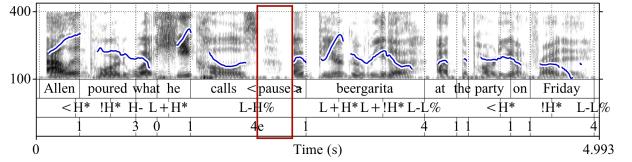


Figure 2: Pitch track of ‘Allen poured what he calls % a beergarita % at the party on Friday.’ (data from SQ Exp.) This example demonstrates the presence of an emphatic juncture preceding the perspective-shifted material in a transparent free relative.

#### 4.1.2. The emphatic juncture in SQ

The Emphatic Juncture is also used to mark the beginning of SQ. This is perhaps unsurprising given that SQ has also been analyzed semantically as a perspective-shifting phenomenon. Of the 61 tokens examined, 60 contained an emphatic juncture marking the start of SQ. Some examples of this can be seen in Figs. 3 and 4 below.

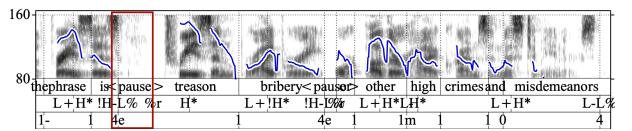


Figure 3: Pitch track of ‘The phrase is % “treason, bribery % and other high crimes and misdemeanors.”’ (data from NPR) This example demonstrates the presence of an emphatic juncture before the quotation.

#### 4.1.3. Why the emphatic juncture?

The Emphatic Juncture consists of an unexpected combination of acoustic features. On the one hand, the boundary tone sequence is a plateau, which indicates an impending continuation and lack of completion [16]. On the other hand, the presence of

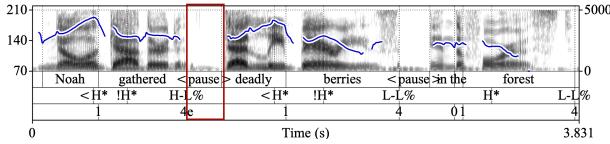


Figure 4: Pitch track of 'Noah gathered % “deadly berries” % in the forest.' (data from SQ Exp.) This example demonstrates the presence of an emphatic juncture before the quotation.

an obligatory pause forces a certain level of discontinuity within the utterance.

These features are naturally in tension and thus form an intriguing pair, but their combination is remarkably sensible for marking the start of SQ. Recall that one of the semantic functions of SQ is to mark an instance of perspective shift. Inserting a large pause is one of the best tools a speaker has to differentiate perspective-shifted material in the midst of an utterance that also contains content the speaker does want to endorse. Meanwhile, the plateau preceding the pause indicates that there is more content following the pause that is related to the material preceding the juncture. By combining both the plateau and the pause, the speaker is cuing they are shifting perspective in the midst of the utterance.

## 4.2. Pitch range reset

### 4.2.1. Pitch range during the quotation

The second critical prosodic feature of SQ is the pitch range reset. In every utterance examined that did not use the lexical marker “quote,” the speaker reset their pitch range at the beginning of the quotation. In fact, this reset could often also be characterized as a pitch range expansion. The following examples in Figs. 5 and 6 demonstrate the pitch range reset associated with SQ. Following MAE.ToBI convention, pitch range reset is annotated as *%r* in the Tones tier.

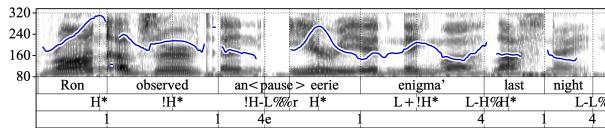


Figure 5: Pitch track of 'Ron observed an % “eerie enigma” % last night.' (data from SQ Exp.) Example demonstrating pitch range reset.

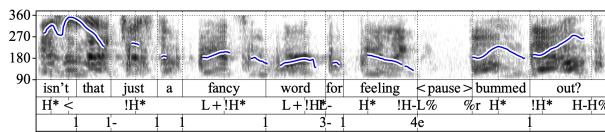


Figure 6: Pitch track of 'Isn't that just a fancy word for feeling % “bummed out?”' (data from "The Office" S3E20 "Safety Training") Example demonstrating pitch range reset.

### 4.2.2. Why the pitch range reset?

Resetting the pitch range is a means for the speaker to increase the prominence of the content of the quotation. It also likely serves as another indicator in addition to the preceding emphatic

juncture that the material within the quotation is distinct from the preceding content at an information structural level. In particular, the SQ content has a source other than the speaker.

### 4.2.3. Pitch range following the quotation

Interestingly, following the end of the SQ, the speaker often returns to their pre-quotation pitch register. In fact, in many cases the material within the quotation can be spliced out to leave a relatively continuous-sounding pitch track. When low-pass filtered, the spliced version sounds like a felicitous English contour. The pair of spectrograms in Fig. 7 demonstrate the original sentence, with SQ around *disastrous design*, above and the spliced version below.

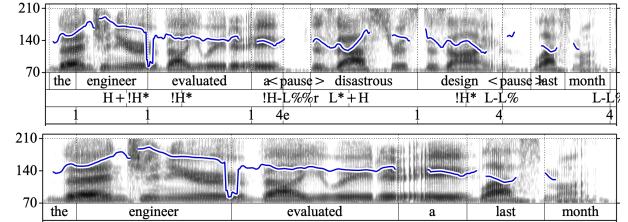


Figure 7: Pitch tracks of 'The engineer evaluated a (% “disastrous design” %) last month.' (data from SQ Exp.) A pair of sentences demonstrating pitch track continuity before and after the quotation. The original sentence is above and the spliced sentence without the object is below.

## 4.3. Quotation final IP break

The final prosodic feature of SQ is an IP break at the right edge of the quoted material. The presence of this break is most apparent in sentence medial instances of SQ, which eliminates the boundary being confounded with sentence final boundary tones. The boundary tone sequence associated with the IP break at the end of the quotation is typically L-L% or L-H%. An example of each can be seen in Figs. 8 and 9 below. Of the 41 SQ tokens uttered sentence medially, 39 included a right edge IP break.

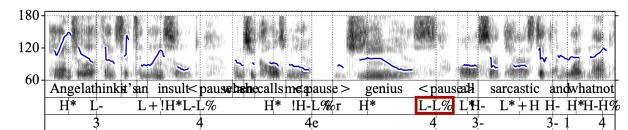


Figure 8: Pitch track of 'Angelathinks it's an insult when she calls me a % “genius” % all sarcastic and whatnot?' (data from "The Accountants" webisode of "The Office") This example shows an IP break at the end of the quotation with an L-L% boundary tone sequence.

Some utterances, like the examples in Figs. 8 and 9, included a notable pause after the SQ, similar to the emphatic juncture at the start of the quotation. This pause seems to be optional, however, as other utterances did not include a quotation-final pause. An example omitting the pause is shown in Fig. 10 below.

### 4.3.1. Why the quotation final IP?

Together, the emphatic juncture at the beginning of the quotation and the IP break at the end form something akin to prosodic

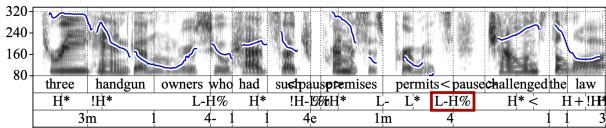


Figure 9: Pitch track of 'Three handgun owners with such % "premises permits" % challenged the law...' (data from NPR) This example shows an IP break at the end of the quotation with an L-H% boundary tone sequence.

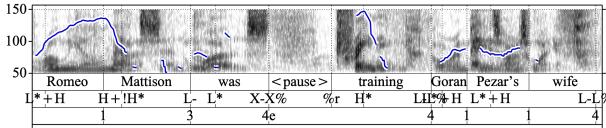


Figure 10: Pitch track of 'Romeo Mattison was % "training" % Goran Pezar's wife.' (data from "Barry" SIE3: "Make the Unsafe Choice") This example shows an IP break at the end of the quotation with no notable pause.

bracketing. These junctures help the speaker communicate to the listener the content they wish to attribute to another source or otherwise demarcate as pragmatically unexpected or out of the ordinary. The IP break at the end of SQ indicates a return to the default speaker-oriented perspective.

## 5. Using “Quote”

In some cases, the speaker can lexically mark the beginning of SQ using “quote.” This was found most often in newscaster speech from the NPR data. There were also a few tokens from various officials testifying in congressional hearings, also heard on NPR. Some examples of utterances with the beginning of SQ preceded by “quote” can be seen in Figs. 11 and 12.

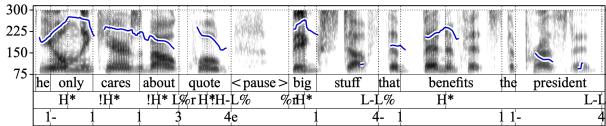


Figure 11: Pitch track of 'He only cares about quote % "big stuff, % that benefits the president.'" (data from NPR) This example shows the use of “quote” by a newscaster.

The most common function of these types of quotes was subclausal direct speech reports, particularly in reporting about politics or witnesses testifying. In these contexts, it is important to accurately capture both the exact words the person being quoted used as well as to make clear that those particular words are not necessarily the beliefs of the individual speaking or their organization.

### 5.1. Prosody of SQ with “quote”

The prosody associated with SQ is remarkably consistent whether or not “quote” is used. When the speaker does use “quote,” however, all three of the required prosodic markers for bare SQ need not be strictly observed. An emphatic juncture occurs directly following “quote” (marked as 4e in Figs. 11 and 12 above). The speaker can optionally employ a pitch range reset, such as in Fig. 11, but it is not obligatory when “quote” is

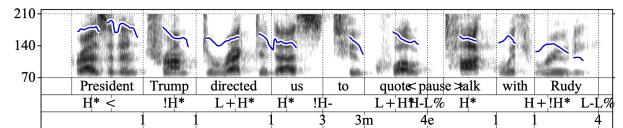


Figure 12: Pitch track of 'President Trump direct us to quote % "talk with Rudy"' (data from NPR) This example shows the use of “quote” by a witness during a congressional hearing.

used, such as in Fig. 12.

Interestingly, despite a lexical quotation end marker (“unquote”) being available, speakers very rarely employ it for SQ. Of the 12 NPR tokens that utilized “quote,” none of them included a paired “unquote.” Unfortunately, none of these “quote” examples occurred sentence medially, so it is not yet clear whether the right edge IP break is obligatory for SQ marked with “quote.”

## 6. Conclusion

In this paper we have explored the prosodic features associated with subclausal quotation in Mainstream American English. Using data from NPR, television shows, and a production experiment with read sentences, we identified three key characteristics of SQ intonation. The beginning of SQ is marked with an emphatic juncture. Second, the quoted material involves a pitch range reset. Finally, the end of the quotation is marked by an L-L% or L-H% IP break. These same characteristics can be used when speaker lexically marks the quotation with “quote,” though using this marker allows for more optionality.

Future research could address the prosody associated with SQs marked with “unquote” when it is employed (though rarely.) Moreover, this work has focused primarily on the production of SQ. It could prove insightful to explore the perception of these prosodic markers as well. For instance, how many prosodic cues are necessary for a listener to perceive SQ? Are the cues gradient or categorical? Is prosody alone sufficient to convey the presence of SQ, or must SQ prosody also be accompanied by multi-modal features such as air quotes and facial expressions?

In the quest to incorporate intonation more rigorously into formal semantics, it may also be helpful to explore the intonation of semantically related phenomena. If intonational commonalities exist, they could provide insight into the formal role of intonation in conveying meaning. Given that both SQ and transparent free relatives mark the beginning of perspective-shifted material intonationally with an emphatic juncture, this sort of inquiry already seems promising.

## 7. References

- [1] E. Kasimir, “Prosodic correlates of subclausal quotation marks,” *ZAS Papers in Linguistics*, vol. 49, pp. 67–77, 2008.
- [2] M. Beckman, J. Hirschberg, and S. Shattuck-Hufnagel, “The original ToBI system and the evolution of the ToBI framework,” in *Prosodic Typology: The Phonology of Intonation and Phrasing, Volume 1*, S. Jun, Ed. Oxford University Press, 2005, pp. 9–54.
- [3] M. Beckman and G. Ayers-Elam, “Guidelines for ToBI labelling (version 3.0),” 1997.
- [4] J. Pierrehumbert, “The phonology and phonetics of english intonation,” PhD Thesis, MIT, 1980.
- [5] M. Beckman and J. Pierrehumbert, “Intonational structure in

- japanese and english," *Phonology Yearbook*, vol. 3, pp. 225–310, 1986.
- [6] H. H. Clark and R. J. Gerrig, "Quotations as demonstrations," *Language*, vol. 66, pp. 764–805, 1990.
  - [7] W. Jansen, M. L. Gregory, and J. M. Brenier, "Prosodic correlates of directly reported speech: Evidence from conversational speech," in *ISCA tutorial and research workshop (ITRW) on prosody in speech recognition and understanding*, 2001.
  - [8] C. Potts, "The dimensions of quotation," *Direct compositionality*, pp. 405–431, 2007.
  - [9] J. Harris and C. Potts, "Perspective-shifting with appositives and expressives," *Linguistics and Philosophy*, vol. 32, pp. 523–552, 2009.
  - [10] B. Geurts and E. Maier, "Quotation in context," *Belgian Journal of Linguistics*, vol. 17, pp. 109–128, 2005.
  - [11] P. Lasersohn, "Expressives, perspective and presupposition," *Theoretical Linguistics*, vol. 33, no. 2, pp. 223–230, 2007.
  - [12] C. S. Smith, *Modes of discourse: The local structure of texts*. Cambridge University Press, 2003, vol. 103.
  - [13] B. Sturman, "The emphatic juncture: A novel use of the IP boundary in English," in *Proceedings of the 19th International Congress of Phonetic Sciences*, S. Calhoun, P. Escudero, M. Tabain, and P. Warren, Eds. Melbourne, Australia: Canberra, Australia: Australasian Speech Science and Technology Association Inc., 2019, pp. 2489–2493.
  - [14] A. Grosu, "'transparent' free relatives as a special instance of 'standard' free relatives," *From NP to DP*, vol. 1, pp. 139–178, 2003.
  - [15] C. Wilder, "Transparent free relatives," *ZAS papers in linguistics*, vol. 10, pp. 191–199, 1998.
  - [16] J. Pierrehumbert and J. Hirschberg, "The meaning of intonational contours in the interpretation of discourse," in *Intentions in Communication*, P. Cohen, Morgan, and Pollack, Eds. MIT Press, 1990, pp. 271–311.